JOURNAL OF MULTIVARIATE ANALYSIS 27, 15–23 (1988)

Joint Asymptotic Distribution of Marginal Quantiles and Quantile Functions in Samples from a Multivariate Population*

G. JOGESH BABU

The Pennsylvania State University

AND

C. RADHAKRISHNA RAO

University of Pittsburgh

Communicated by the Editors

The joint asymptotic distributions of the marginal quantiles and quantile functions in samples from a p-variate population are derived. Of particular interest is the joint asymptotic distribution of the marginal sample medians, on the basis of which tests of significance for population medians are developed. Methods of estimating unknown nuisance parameters are discussed. The approach is completely nonparametric.

1. INTRODUCTION

Let $X=(x_1, ..., x_p)$ be a random vector with joint d.f. (distribution function) F, ith marginal d.f. F_i, (i, j)th marginal d.f. F_{ij} and ith marginal density function f_i. We denote the ith marginal quantile function by

$$\xi_i(q)=F_i^{-1}(q)=\inf\{x: F_i(x)\geqslant q\}, \qquad 0<q<1 \tag{1.1}$$

and, for convenience, a specific quantile say the q_ith of F_i by

$$\theta_i=\xi_i(q_i). \tag{1.2}$$

Received November 11, 1987.

AMS 1980 subject classifications: 62H15, 62G30.

Key words and phrases: asymptotic tests, bootstrap, quantiles, quantile processes, tests based on medians.

* Research sponsored by the Air Force Office of Scientific Research under Contract F49620-85-C-0008. The U.S. Government's right to retain a nonexclusive royalty-free license in and to the copyright covering this paper, for governmental purposes, is acknowledged.

0047-259X/88 $3.00

Further, let

$$\eta_{ij}(q, r) = F_{ij}(\xi_i(q), \xi_j(r)) \tag{1.3}$$

and denote for given q_i and q_j,

$$\sigma_{ij} = \eta_{ij}(q_i, q_j) - q_i q_j = F_{ij}(\theta_i, \theta_j) - q_i q_j. \tag{1.4}$$

The parameters (1.1)–(1.4) defined above refer to the d.f. of X.

Now let

$$X_i = (x_{1i}, ..., x_{pi}), \qquad i, ..., n \tag{1.5}$$

be n independent copies of X and denote the empirical d.f. of $\{X_i, i = 1, ..., n\}$ by $F^{(n)}$ and the corresponding ith and (i, j)th marginal distributions by $F_i^{(n)}$ and $F_{ij}^{(n)}$, respectively. We denote the quantities (1.1)–(1.4) defined in terms of $F^{(n)}$, $F_i^{(n)}$, and $F_{ij}^{(n)}$ by

$$\xi_i^{(n)}(q), \qquad \theta_i^{(n)}, \qquad \text{and} \qquad \sigma_{ij}^{(n)} \tag{1.6}$$

or simply as

$$\hat{\xi}_i(q), \qquad \hat{\theta}_i, \qquad \text{and} \qquad \hat{\sigma}_{ij} \tag{1.7}$$

as estimates of $\xi_i(q)$, θ_i, and σ_{ij}, respectively.

In this paper, we derive the asymptotic distribution of

$$\hat{\boldsymbol{\theta}}' = (\hat{\theta}_1, ..., \hat{\theta}_p) = (\hat{\xi}_1(q_1), ..., \hat{\xi}_p(q_p)) \tag{1.8}$$

for given $q_1, ..., q_p$ and also the joint distribution of the marginal quantile processes

$$\hat{\xi}_i(q), \qquad 0 < q < 1, i = 1, ..., p. \tag{1.9}$$

The asymptotic distributions of the empirical quantle process (Csörgö and Révész [6]) and of a fixed set of specified quantiles (Mosteller [11]) in one dimension are well known.

Of particular interest is the joint asymptotic distribution of the marginal sample medians

$$(\hat{\xi}_1(\tfrac{1}{2}), ..., \hat{\xi}_p(\tfrac{1}{2})) \tag{1.10}$$

using which we develop tests of significance for the population medians analogous to tests for the means in the multivariate case (see Rao [12, pp. 543–573]). An early work on the joint asymptotic distribution of the sample medians is due to Mood [10]; see also Kuan and Ali [8], where they assume the existence of the density function for the vector variable X.

We obtain the dsitribution in the general case in a form convenient for practical applications.

2. Distribution of the Marginal Sample Quantiles

We prove the following theorem concerning the joint asymptotic distribution of

$$(\hat{\theta}_1, ..., \hat{\theta}_p) = (\hat{\xi}_1(q_1), ..., \hat{\xi}_p(q_p)), \tag{2.1}$$

the sample q_1th, ..., q_pth quantiles of the marginal empirical distributions of $x_1, ..., x_p$, respectively.

Theorem 2.1. *Let F_i be continuously twice differentiable in a neighborhood of θ_i and $\delta_i = f_i(\xi_i(q_i)) = f_i(\theta_i) > 0$, $i = 1, ..., p$, where f_i denotes the derivative of F_i. Then the asymptotic distribution of*

$$y_n = \sqrt{n}(\hat{\theta}_1 - \theta_1, ..., \hat{\theta}_p - \theta_p) \tag{2.2}$$

is p-variate normal with mean vector zero, and variance-covariance matrix

$$\Sigma = \begin{pmatrix} \dfrac{q_1(1-q_1)}{\delta_1^2} & \dfrac{\sigma_{12}}{\delta_1 \delta_2} & \cdots & \dfrac{\sigma_{1p}}{\delta_1 \delta_p} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\sigma_{p1}}{\delta_p \delta_1} & \dfrac{\sigma_{p2}}{\delta_p \delta_2} & \cdots & \dfrac{q_p(1-q_p)}{\delta_p^2} \end{pmatrix}, \tag{2.3}$$

where σ_{ij} are as defined in (1.4).

Proof. By Bahadur's representation of the sample quantiles (see Bahadur [4]),

$$(\log n)^{-1} n^{3/4} \, |(\hat{\theta}_i - \theta_i) - \delta_i^{-1}(r_i - q_i)| \xrightarrow{P} 0, \qquad i = 1, ..., p, \tag{2.4}$$

where $r_i = F_i^{(n)}(\theta_i)$. Then, it follows that

$$y_n = \sqrt{n}(\hat{\theta}_1 - \theta_1, ..., \hat{\theta}_p - \theta_p) \tag{2.5}$$

and

$$z_n = \sqrt{n}(\delta_1^{-1}(r_1 - q_1), ..., \delta_p^{-1}(r_p - q_p)) \tag{2.6}$$

have the same asymptotic distribution. By the multivariate central limit theorem, z_n weakly converges to a p-variate normal distribution with mean vector zero and covariance matrix as given in (2.3). This proves Theorem 2.1.

For practical applications we need a consistent estimate of Σ as defined in (2.3). There are two sets of unknown $\{\sigma_{ij}\}$ and $\{\delta_i^{-1}\}$ in Σ. A consistent estimate of σ_{ij} is provided by $\hat{\sigma}_{ij}$ as shown in Theorem 2.2.

THEOREM 2.2. *Let F_{ij} be continuous at $(\theta_i, \theta_j) = (\xi_i(q_i), \xi_j(q_j))$. Then*

$$\hat{\sigma}_{ij} = F_{ij}^{(n)}(\xi_i^{(n)}(q_i), \xi_j^{(n)}(q_j)) = F_{ij}^{(n)}(\hat{\theta}_i, \hat{\theta}_j) \to \sigma_{ij} = F_{ij}(\theta_i, \theta_j) \quad \text{a.e. as} \quad n \to \infty. \tag{2.7}$$

Proof.

$$|F_{ij}(\theta_i, \theta_j) - F_{ij}^{(n)}(\hat{\theta}_i, \hat{\theta}_j)| \leqslant |F_{ij}(\theta_i, \theta_j) - F_{ij}(\hat{\theta}_i, \hat{\theta}_j)| + \sup_{x, y} |F_{ij}(x, y) - F_{ij}^{(n)}(x, y)|. \tag{2.8}$$

Since F_{ij} is continuous at (θ_i, θ_j) and

$$\sup |F_{ij}(x, y) - F_{ij}^{(n)}(x, y)| \to 0 \qquad \text{a.e.} \tag{2.9}$$

it follows that the expression on the left-hand side of (2.8) $\to 0$ a.e. which establishes the result (2.7) of Theorem 2.2. Equation (2.9) is a consequence of Theorem 7.2 of Rao [13].

The result (2.7) implies that σ_{ij} in (2.3) can be consistently estimated by its sample equivalent $\hat{\sigma}_{ij}$.

There exist several methods for the estimation of δ_i (see Krieger and Pickards, III [7] and the references therein). Recently, a consistent and efficient estimator of δ_i^{-1} based on a sample of size n has been proposed by Bahu [2] under the assumption that f_i is continuously differentiable at $\xi_i(q_i)$. There is a possibility of this estimate taking negative values, and when this happens some modification of the estimate may have to be made. Using consistent estimates of $\hat{\sigma}_{ij}$ and δ_i^{-1}, a consistent estimate of $\sigma_{ij}/\delta_i\delta_j$, the (i, j)th element of Σ, can be obtained as $\hat{\sigma}_{ij}/\hat{\delta}_i\hat{\delta}_j$.

Another possibility is to obtain a direct estimate of $\sigma_{ij}/\delta_i\delta_j$ by the bootstrap method

$$\hat{\sigma}_{ij}/\hat{\delta}_i\hat{\delta}_j = E^*[n(\theta_i^* - \hat{\theta}_i)(\theta_j^* - \hat{\theta}_j)] \tag{2.10}$$

where E^* is the expectation under the bootstrap distribution function. The consistency of the estimator (2.10) can be proved on the same lines as those given by Babu [3] for the bootstrap estimate of the variance of the sample median.

3. Tests of Significance Based on Medians

Let

$$\hat{\boldsymbol{\theta}}_i' = (\hat{\theta}_{1i}, ..., \hat{\theta}_{pi}), \hat{\Sigma}_i \tag{3.1}$$

be the marginal sample medians and an estimate of Σ (as defined in (2.3)) obtained from a sample of size n_i from a p-variate population Π_i, $i = 1, ..., k$. Further let $\boldsymbol{\theta}_i = (\theta_{1i}, ..., \theta_{pi})$ be the true value of the marginal medians for Π_i. To test the hypothesis

$$\boldsymbol{\theta}_1 = \cdots = \boldsymbol{\theta}_p \tag{3.2}$$

we can use the statistic

$$\chi^2 = \text{trace}\left[\sum_{i=1}^{k} n_i \Sigma_i^{-1} \hat{\boldsymbol{\theta}}_i \hat{\boldsymbol{\theta}}_i' - \left(\sum_{i=1}^{k} n_i \Sigma_i^{-1}\right) \bar{\boldsymbol{\theta}} \bar{\boldsymbol{\theta}}'\right], \tag{3.3}$$

where

$$\bar{\boldsymbol{\theta}} = \left(\sum_{i=1}^{k} n_i \Sigma_i^{-1}\right)^{-1} \sum_{i=1}^{k} n_i \Sigma_i^{-1} \hat{\boldsymbol{\theta}}_i \tag{3.4}$$

as chi-square on $p(k-1)$ degrees of freedom, provided the individual sample sizes $n_1, ..., n_k$ are large.

In cases where a common Σ for the k populations can be assumed, we have the problem of estimating Σ from the combined sample. For this purpose we consider the residual vectors by replacing each observed vector by its difference from the sample median vector computed from the sample to which the observed vector belongs. There are altogether $n = (n_1 + \cdots + n_k)$ residual vectors, arising out of the k different samples, from which we construct a p-dimensional empirical distribution function E with the marginal medians as zeros. Then σ_{ij} can be estimated from E_{ij}, the (i, j)th marginal d.f. of E as indicated in (2.7) and δ_i from E_i, the ith marginal d.f. of E using any of the methods described at the end of Section 2. If we denote a common estimate of Σ by $\hat{\Sigma}$, then we can develop tests of significance concerning the structure of the median vectors $\boldsymbol{\theta}_i$, $i = 1, ..., k$, as in the case of mean values (see Rao [12, p. 556]). For this purpose we compute the "between populations" matrix

$$S = \sum_{i=1}^{k} n_i \hat{\boldsymbol{\theta}}_i \hat{\boldsymbol{\theta}}_i' - n \bar{\boldsymbol{\theta}} \bar{\boldsymbol{\theta}}' \tag{3.5}$$

where $n\bar{\boldsymbol{\theta}} = n_1 \hat{\boldsymbol{\theta}}_1 + \cdots + n_k \hat{\boldsymbol{\theta}}_k$, and set up the determinental equation

$$|S - \lambda \hat{\Sigma}| = 0. \tag{3.6}$$

The roots of Eq. (3.6) can be used as in the table on p. 558 of Rao [12] to test the dimensionality of the configuration of median values.

4. Joint Distribution of the Marginal Quantile Processes

In Section 2 of the paper, we derived the joint asymptotic distribution of specified marginal quantiles. We now derive the weak limits of the entire marginal quantile processes after suitable scaling. More specifically we consider the processes $\{Z_n\}$ indexed by $(q_1, ..., q_p) \in (0, 1)^p$, where

$$Z_n(q_1, ..., q_p) = \sqrt{n}\,[f_1(\xi_1(q_1))(\xi_1^{(n)}(q_1) - \xi_1(q_1)), ..., f_p(\xi_p(q_p))(\xi_p^{(n)}(q_p) - \xi_p(q_p))]. \tag{4.1}$$

We first simplify the problem using the following result which is essentially a restatement of Theorem 5.2.2 of Csörgö and Révész [6].

Theorem 4.1. *Suppose that for $i = 1, ..., p$, the marginal d.f. F_i is twice differentiable on (a_i, b_i), where*

$$-\infty \leqslant a_i = \sup\{x : F_i(x) = 0\}$$
$$\infty \geqslant b_i = \inf\{x : F_i(x) = 1\}$$

and $F_i' = f_i \neq 0$ on $(a_i b_i)$. Further assume that

$$\max_i \sup_{a_i < x < b_i} F_i(x)[1 - F_i(x)] \frac{|f_i'(x)|}{f_i^2(x)} < \infty$$

and f_i is non-decreasing (non-increasing) on an interval to the right of a_i (to the left of b_i). Let

$$Y_n^*(q_1, ..., q_p) = \sqrt{n}(V_1^{(n)}(q_1) - q_1, ..., V_p^{(n)}(q_p) - q_p),$$

where $V_i^{(n)}$ is the empirical d.f. of the uniform variables

$$u_{ij} = F_i(x_{ij}), \qquad j = 1, ..., n.$$

Then

$$\sup_{\mathbf{q} \in (0,1)^p} \| Y_n^*(\mathbf{q}) - Z_n(\mathbf{q}) \| \to 0 \qquad a.e. \tag{4.2}$$

Hence $\{Y_n^*\}$ and $\{Z_n\}$ have the same limit.

Note that the marginals of $\{Y_n^*\}$ converge weakly to a Brownian bridge on $C[0, 1]$ (see Billingsley [5, p. 105]). Since the paths of the limiting process are continuous, we define a new process Y_n close to Y_n^* as follows. Let $D_1^{(n)}(t)$ be, as a function of $t \in [0, 1]$, the d.f. corresponding to a uniform distribution of mass $(n+1)^{-1}$ over each of the $(n+1)$ intervals $[d_{j-1}, d_j]$, $j = 1, ..., n+1$, where $d_0 = 0$, $d_{n+1} = 1$, and $d_1, ..., d_n$ are the values of $u_{i1}, ..., u_{in}$ arranged in increasing order. Clearly

$$|V_i^{(n)}(t) - D_i^{(n)}| \leqslant \frac{1}{n}, \qquad 0 \leqslant t \leqslant 1 \text{ a.e.}$$

So if

$$Y_n(\mathbf{q}) = \sqrt{n}(D_1^{(n)}(q_1) - q_1, ..., D_p^{(n)}(q_p) - q_p)$$

then

$$\| Y_n(\mathbf{q}) - Y_n^*(\mathbf{q})\| \leqslant n^{-1/2} \qquad \forall \mathbf{q} \in [0, 1]^p \text{ a.e.}$$

As a consequence, $\{Y_n\}$ and $\{Z_n\}$ have the same weak limits and the marginals of Y_n are continuous functions. Note that

$$Y_n \in B = \{h\colon h(\mathbf{q}) = (h_1(q_1), ..., h_p(q_p)), h_i \text{ is a continuous function on } [0, 1], i = 1, ..., p\}.$$

Clearly B is a separable closed linear subspace of the Banach space C_p of continuous functions on $[0, 1]^p$ into $\mathbb{R}^p$.

We shall show that $\{Y_n\}$ converges weakly to a Gaussian measure on B. A probability measure μ on B is called Gaussian if for every $H \in B^*$, the space of real continuous linear functionals on B, μH^{-1} is Gaussian on the line (see Aranjo and Giné [1, pp. 140–142, 28, and problem 2 on p. 33]).

To characterize B^*, let H be a real continuous linear functional on B. Then

$$\begin{aligned} H(h_1, ..., h_p) &= H(h_1, 0, ..., 0) + \cdots + H(0, 0, ..., h_p) \\ &= H_1(h_1) + \cdots + H_p(h_p), \qquad \text{say.} \end{aligned} \tag{4.3}$$

The zeroes in the first line of (4.3) refer to the zero function. Clearly, each H_i is a real continuous linear functional on $C[0, 1]$. It then follows that B^* is the k-fold direct sum of the dual space C^* of $C[0, 1]$. By Riesz's representation theorem, for any $L \in C^*$, there exists a signed measure ν on $[0, 1]$ such that

$$L(f) = \int_0^1 f(x)\, d\nu(x)$$

for any $f \in C[0, 1]$ (see Dunford and Schwartz [9]). Thus for every $H \in B^*$, there exist signed measures $\nu_1, ..., \nu_p$ on $[0, 1]$ such that for $f = (f_1, ..., f_p) \in B$,

$$H(f) = \sum_{i=1}^{p} \int_0^1 f_i(x)\, d\nu_i(x).$$

Now let

$$A = \left\{ \sum_{j=1}^{r} \alpha_j \varepsilon_{x_j} : 0 \leqslant x_j \leqslant 1,\, x_j,\, \alpha_j \text{ rational},\, j = 1, ..., r,\, r = 1, 2, ... \right\},$$

where ε_x is the probability measure putting all its mass at x. It is easily seen that A is dense in C^* and is countable. We now state the main result.

THEOREM 4.2. *$\{Y_n\}$ converges weakly to a Gaussian random element $W = (W_1, ..., W_k)$ in B, where W_i is a Brownian bridge for each i and*

$$E(W_i(t)\, W_j(s)) = P(F_i(x_{i1}) \leqslant t,\, F_j(x_{j1}) \leqslant s) - ts \tag{4.4}$$

for all i, j and $0 \leqslant t,\ s \leqslant 1$.

Proof. Since $\{\sqrt{n}(D_i^{(n)}(t) - t) : 0 \leqslant t \leqslant 1\}$ is tight for each i in $C[0, 1]$, it follows that $\{Y_n\}$ is tight in B. Since A is dense in C^*, in order to show that $\{Y_n\}$ has a weak limit it is enough to show that for any $q_{11}, ..., q_{1r}, ..., q_{p1}, ..., q_{pr}$ in $[0, 1]$ and α_{ij} real

$$\sum_{i=1}^{p} \sum_{j=1}^{r} \alpha_{ij} \sqrt{n} (D_i^{(n)}(q_{ij}) - q_{ij})$$

converges weakly. This holds because of the central limit theorem and the fact that

$$\sup_{0 \leqslant t \leqslant 1} |V_i^{(n)}(t) - D_i^{(n)}(t)| \leqslant \frac{1}{n} \qquad \text{a.e.}$$

To complete the proof it is enough to show the existence of W satisfying (4.4).

Since $\{Y_n\}$ is tight, there exists a random element Y on B and a subsequence $\{Y_{n'}\}$ such that $Y_{n'}$ converges weakly to $Y = (Y^{(1)}, ..., Y^{(p)})$. Further, from the above arguments

$$\sum_{i=1}^{p} \sum_{j=1}^{r} \alpha_{ij} Y^{(i)}(q_{ij}) \qquad \text{and} \qquad \sum_{i=1}^{p} \sum_{j=1}^{r} \alpha_{ij} W_i(q_{ij})$$

have the same distribution as that of normal random variables. So it follows that Y satisfies the properties of W mentioned in (4.4) and Y is Gaussian. Thus Y_n converges weakly to W, and in view of Theorem 4.1, $\{Z_n\}$ converges to W.

References

[1] Aranjo, A., and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables.* Wiley, New York.

[2] Babu, G. J. (1986). Efficient estimation of the reciprocal of the density quantile function at a point. *Statist. Probab. Lett.* **4** 133–139.

[3] Babu, G. J. (1986). A note on bootstrapping the variance of sample quantiles. *Ann. Inst. Statist. Math.* **38** 439–443.

[4] Bahadur, R. R. (1966). A note on quantiles in large samples. *Ann Math. Statist.* **37** 577–580.

[5] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[6] Csörgő, M., and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

[7] Krieger, A. M., and Pickands, III, J. (1981). Weak convergence and efficient density estimation at a point. *Ann. Statist.* **9** 1066–1078.

[8] Kuan, K. S., and Ali, M. M. (1980). Asymptotic distribution of quantiles from a multivariate distribution. *Multivariate Statistical Analysis, Proceedings, Conference at Dalhousie University, Halifax, Nova Scotia, 1979*, pp. 109–120, North-Holland, Amsterdam.

[9] Dunford, N., and Schwartz, J. T. (1958). *Linear Operators*, Part I. Interscience, New York.

[10] Mood, A. M. (1941). On the joint distribution of the medians in samples from a multivariate population. *Ann. Math. Statist.* **12** 268–278.

[11] Mosteller, F. (1946). On some useful "inefficient statistics." *Ann. Math. Statist.* **17** 377–408.

[12] Rao, C. R. (1973). *Linear Statistical Inference and Its Application*, 2nd ed., Wiley, New York.

[13] Rao, R. R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Mat. Statist.* **33** 659–680.